

# 李润辉

联系电话 (+852) 9517 4226

电子邮件 lrhdiy@gmail.com

地址 香港沙田香港科学园核心大楼 1 座 335-337 室

个人主页 rhli.github.io

年龄 28

## 简介

专注分布式存储系统性能及可靠性，五年以上分布式存储系统开发经验，在多个项目中修改 Hadoop 源码，了解并使用内存数据库（如 Memcached，Redis）

编程语言: C++, C, Java, Python

## 工作经历

2016–Now 研究员, 华为未来网络理论实验室.

2015–2016 博士后研究员, 香港中文大学.

## 教育经历

2011–2015 博士, 计算机科学与工程学, 香港中文大学.

导师: Patrick P. C. Lee

毕业论文: 构建基于纠删码的高效和高可靠性文件系统的研究

2007–2011 工学学士, 计算机科学与技术, 中国科学技术大学.

## 研究兴趣

大数据存储和处理, 内存数据库

## 发表论文

2017 Qun Huang, Xin Jin, Patrick P. C. Lee, Runhui Li, Lu Tang, Yi-Chao Chen, Gong Zhang

**Repair Pipelining for Erasure-Coded Storage.**

The 2017 USENIX Annual Technical Conference (ATC'17), Santa Clara, California, July 2017

2017 Runhui Li, Xiaolu Li, Patrick P. C. Lee and Qun Huang

**Repair Pipelining for Erasure-Coded Storage.**

香港科学园核心大楼 1 座 335-337 室 – 沙田 – 香港

☎ (852) 9517 4226 • ✉ lrhdiy@gmail.com • 🌐 rhli.github.io

- The 2017 USENIX Annual Technical Conference (ATC'17), Santa Clara, California, July 2017
- 2017 Eman Ramadan, Arvind Narayanan, Zhi-Li Zhang, Runhui Li and Gong Zhang  
**BIG Cache Abstraction for Cache Networks.**  
The 37th IEEE International Conference on Distributed Computing Systems (ICDCS'17), Atlanta, Georgia, June 2017
- 2017 Runhui Li, Yuchong Hu and Patrick P. C. Lee  
**Enabling Efficient and Reliable Transition from Replication to Erasure Coding for Clustered File Systems.**  
Transaction on Parallel and Distributed Systems (TPDS) (to appear)
- 2015 Mingqiang Li, Runhui Li and Patrick P. C. Lee  
**Relieving Both Storage and Recovery Burdens in Big Data Cluster with R-STAIR Code.**  
2015 USENIX Annual Technical Conference (ATC'15) (Poster presentation), Santa Clara, CA, July 2015
- 2015 Runhui Li, Jian Lin and Patrick P. C. Lee  
**Enabling Concurrent Failure Recovery for Regenerating-Coding-Based Storage Systems: From Theory to Practice.**  
IEEE Transaction on Computers (TC), 64(7), pp. 1898-1911, July 2015
- 2015 Runhui Li, Yuchong Hu and Patrick P. C. Lee  
**Enabling Efficient and Reliable Transition from Replication to Erasure Coding for Clustered File Systems.**  
Proceedings of 45th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN'15) (Regular paper), Rio de Janeiro, Brazil, June 2015  
(AR:  $50/229 = 21.8\%$ )
- 2015 Runhui Li and Patrick P. C. Lee  
**Making MapReduce Scheduling Effective in Erasure-Coded Storage Clusters.**  
Proceedings of the 21st IEEE International Workshop on Local and Metropolitan Area Networks (LANMAN'15) (Invited paper), Beijing, China, April 2015
- 2014 Runhui Li, Patrick P. C. Lee and Yuchong Hu  
**Degraded-First Scheduling for MapReduce in Erasure-Coded Storage Clusters.**

- Proceedings of 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN'14) (Regular paper), Atlanta, Georgia, June 2014  
(AR:  $56/181 = 30.9\%$ )
- 2014 Silei Xu, Runhui Li, Patrick P. C. Lee, Yunfeng Zhu, Liping Xiang, Yinlong Xu and John C. S. Lui  
**Single Disk Failure Recovery for X-Code-Based Parallel Storage Systems.**  
IEEE Transaction on Computers (TC), 63(4), pp. 995-1007, April 2014
- 2013 Runhui Li, Jian Lin and Patrick P. C. Lee  
**CORE: Augmenting Regenerating-Coding-Based Recovery for Single and Concurrent Failures in Distributed Storage Systems.**  
Proceedings of 29th IEEE Conference on Massive Data Storage (MSST'13) (Short paper), Long Beach, CA, May 2013  
(AR:  $(14 + 15)/109 = 26.6\%$ )
- 2011 Liping Xiang, Yinlong Xu, John Lui, Qian Chang, Yubiao Pan and Runhui Li  
**A Hybrid Approach to Failed Disk Recovery Using RAID-6 Codes: Algorithms and Performance Evaluation.**  
ACM Transaction on Storage, 7(3):11, Oct 2011

---

## 项目和开源软件

- 2016 提升分布式文件系统修复性能的 **Repair Pipelining** 中间件  
本人职责: 全部设计与实现  
简介: Repair pipelining 是用以提升基于纠删码分布式文件系统 (例如, HDFS, QFS 等) 修复操作的中间件。通过优化网络资源使用, 将修复时间降低多达 **90%** 以上。
- 2016 **ADN (Application-Driven Network)** 原型系统  
本人职责: 主导设计实现控制面  
简介: ADN 通过动态分配网络资源以及为不同应用适配不同网络协议栈, 以提升不同网络服务的 QoS。
- 2015 面向修复的 **R-STAIR** 编码  
本人职责: 主导实现  
简介: R-STAIR 码是一种新型纠删码, 能兼顾高性能的单节点修复性能和多节点错误的容错性。我们在 Hadoop 中实现了 R-STAIR 码。相比传统纠删码, R-STAIR 码能带来 **3.6×** 的单节点修复性能提升。

项目网址: <http://ansrlab.cse.cuhk.edu.hk/software/rstair/>

香港科学园核心大楼 1 座 335-337 室 - 沙田 - 香港

☎ (852) 9517 4226 • ✉ [lrhdiy@gmail.com](mailto:lrhdiy@gmail.com) • 📄 [rhli.github.io](https://github.com/rhli)

**GitHub 地址:** <https://github.com/rhli/hadoop-rstair/>

2015 **EAR: 编码感知的数据复制方案**

**本人职责:** 全部设计与实现

**简介:** 现代的文件系统采用复制和纠删码来分别保护刚写入文件系统的热数据和较少被访问的冷数据。EAR 提供了高效的冷热数据转换策略, 并且提高了转换过程的可靠性。我们在 Hadoop 中实现了 EAR。相比传统的随机复制策略, EAR 将数据由复制转换为纠删码的效率提升了 **120%**。

**项目网址:** <http://ansrlab.cse.cuhk.edu.hk/software/ear/>

**GitHub 地址:** <https://github.com/rhli/hadoop-EAR/>

2014 **面向纠删码存储的 Degraded-First MapReduce 调度算法**

**本人职责:** 全部设计与实现

**简介:** Degraded-First 调度算法, 改善系统中有节点失效时 MapReduce 性能。我们在 Hadoop 中实现了 Degraded-First 算法, 在实际部署中, 可以减少高达 **48%** 的 MapReduce 运行时间。

**项目网址:** <http://ansrlab.cse.cuhk.edu.hk/software/dfs/>

**GitHub 地址:** <https://github.com/rhli/Degraded-First-Scheduler/>

2013 **CORE: 基于再生码的并发故障修复方案**

**本人职责:** 主导设计与实现

**简介:** CORE 扩展现有的再生码, 使之在多节点修复过程中能减少修复开销。我们将 CORE 整合入 Hadoop, 并且进行了验证性实验。相比于传统纠删码, CORE 能带来 **2.33×** 的修复性能提升。

**项目网址:** <http://ansrlab.cse.cuhk.edu.hk/software/core/>

2011 **NCFS: 基于网络编码的分布式文件系统**

**本人职责:** 文件系统层的实现

**简介:** NCFS 是一个验证性的基于网络编码的分布式文件系统。NCFS 是一个基于代理的多节点存储系统。

**项目地址:** <http://ansrlab.cse.cuhk.edu.hk/software/ncfs/>

---

## 获奖经历

2015 The 45th IEEE/IFIP DSN Conference Student Travel Grant

2014 The 44th IEEE/IFIP DSN Conference Student Travel Grant

2011–2015 香港中文大学研究生奖学金

2011 中国科学技术大学优秀本科生毕业论文

2010 花旗奖学金

香港科学园核心大楼 1 座 335-337 室 - 沙田 - 香港

☎ (852) 9517 4226 • ✉ [lrhdiy@gmail.com](mailto:lrhdiy@gmail.com) • 🌐 [rhli.github.io](https://rhli.github.io)

- 2010 中国科学技术大学优秀大学生研究计划
- 2009 中国科学技术大学优秀学生奖学金
- 2008 中国科学技术大学优秀学生奖学金
- 2007 中国科学技术大学优秀新生奖学金